
ELSEVIER
Volume 216, Issue 1, 1 January 2012
ISSN 0377-2217
EUROPEAN JOURNAL OF
OPERATIONAL RESEARCH
Available online at www.sciencedirect.com
SciVerse ScienceDirect
www.elsevier.com/locate/ejor

European Journal of Operational Research 216 (2012) 26–32

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Continuous Optimization

An extension of proximal methods for quasiconvex minimization on the nonnegative orthant

E.A. Papa Quiroz *, P. Roberto Oliveira

Universidad Nacional del Callao, Perú
Universidad Nacional Mayor de San Marcos, Perú
Department of Systems Engineering and Computer Sciences, PESC-COPPE, Federal University of Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 13 July 2010
Accepted 9 July 2011
Available online 23 July 2011

Keywords:
Proximal point methods
Quasiconvex functions
Nonnegative orthant
Proximal distances

ABSTRACT

In this paper we propose an extension of proximal methods to solve minimization problems with quasiconvex objective functions on the nonnegative orthant. Assuming that the function is bounded from below and lower semicontinuous and using a general proximal distance, it is proved that the iterations given by our algorithm are well defined and stay in the positive orthant. If the objective function is quasiconvex we obtain the convergence of the iterates to a certain set which contains the set of optimal solutions and convergence to a KKT point if the function is continuously differentiable and the proximal parameters are bounded. Furthermore, we introduce a sufficient condition on the proximal distance such that the sequence converges to an optimal solution of the problem.

1. Introduction

The class of interior proximal point methods to solve a function f on the nonnegative orthant $\mathbb{R}^n_+$, usually generates a sequence $\{x^k\}$ given by $x^0 \in \mathbb{R}^n_{++}$ (interior of $\mathbb{R}^n_+$) and

$$x^k \in \arg\min_{x \geqslant 0}\{f(x) + \lambda_k d(x, x^{k-1})\}, \tag{1.1}$$

where $x \geqslant 0$ means that $x_i \geqslant 0$, $\forall i = 1,\ldots,n$, and d is a certain function satisfying some desirable properties. For example we can define a like-distance function d to force the iterates x^k to stay in $\mathbb{R}^n_{++}$. Some examples of d based on the literature are the class of Bregman, φ-divergence and second order homogeneous distances (Auslender et al., 1999; Iusem and Teboulle, 1995; Iusem et al., 1994; Teboulle, 1992, 1997; Kiwiel, 1997).

Assuming that f is a proper, lower semicontinuous and convex function, $\mathrm{dom} f \cap \mathbb{R}^n_{++} \neq \emptyset$, and λ_k satisfies $\sum_{k=1}^{+\infty}(1/\lambda_k) = +\infty$, and using a class of proximal distances which contain the three classes mentioned above, Auslender and Teboulle (2006) proved that $\lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \geqslant 0\}$. Furthermore, if the optimal set is nonempty, then sequence $\{x^k\}$ converges to an optimal solution of the problem.

On the other hand, several applications in diverse Science and Engineering areas are sufficient motivation to work with nonconvex objective functions and proximal point methods, see for example, Attouch and Bolte (2009), Goudou and Munier (2009) and Souza et al. (2010). In particular the class of quasiconvex minimization problems has been receiving attention from many researches due to the broad range of applications in location theory (Gromicho, 1998), control theory (Barron and Liu, 1997), and specially in economic theory (Takayama, 1995).

For quasiconvex minimization on the nonnegative orthant there are some recent works in the literature. Attouch and Teboulle (2004), with a regularized Lotka–Volterra dynamical system, have proved the convergence of the continuous method to a point which belongs to certain set which contains the set of optimal points; see also Alvarez et al. (2004), that treats a general class of dynamical systems that includes the one of Attouch and Teboulle (2004). Cunha et al. (2010) and Chen and Pan (2008), with a particular ϕ-divergence distance, have proved the full convergence of the proximal method to the KKT-point of the problem when parameter λ_k is bounded and convergence to an optimal solution when $\lambda_k \to 0$. Pan and Chen (2007), with the second-order homogeneous distance, which includes the Logarithmic–Quadratic proximal point method, and Souza et al. (2010) with a class of separated Bregman distances, have proved the same convergence result of Cunha et al. (2010) and Chen and Pan (2008).

In this paper we are interested in extending the global convergence of a large class of proximal point methods to minimize quasiconvex functions constrained on the nonnegative orthant.

The main difficulty we observed in extending the proximal method for nonconvex function, which was not observed by

* Corresponding author at: Universidad Nacional del Callao, Brazil. Tel.: +55 1 5685537; mob.: +55 1 992332681.
E-mail addresses: erikpapa@gmail.com (E.A. Papa Quiroz), poliveir@cos.ufrj.br (P. Roberto Oliveira).

0377-2217/$ - see front matter Crown Copyright © 2011 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.07.019

previous works, is that due to the nonconvexity of f the subproblems of (1.1) may not be convex and thus, from a practical point of view, we may obtain that minimization subproblems may be as hard to solve globally as the original one due to the existence of multiple isolated local minimizers.

To solve this disadvantage in this paper we propose the following extension for the proximal method: given $x^{k-1} \in \mathbb{R}^n_{++}$, find $x^k \in \mathbb{R}^n_+$ such that

$$0 \in \hat{\partial}(f(.) + \lambda_k d(., x^{k-1}))(x^k), \tag{1.2}$$

where $\hat{\partial}$ is the Fréchet subdifferential and d is a proximal distance, see Sections 2 and 3 respectively. Observe that this condition is more practical than (1.1) where a global minimum point is required in each iteration and therefore more practical than the works of Cunha et al. (2010), Chen and Pan (2008), Souza et al. (2010) and Pan and Chen (2007). Of course both (1.1) and (1.2) are equivalent when f is convex, see Rockafellar and Wets (1998, Proposition 8.12), but in the quasiconvex case these iterative schemes are quite different in nature. Convex problems can be addressed by conventional local algorithms since all critical points are global minimizers. The regularized subproblem being strongly convex when the objective function is convex, we expect the local algorithms to perform efficient enough to find a good approximate solution at reasonable execution time. Suppose now that f is quasiconvex, if the regularized subproblem turned to be quasiconvex, similar considerations as in the convex case would hold true: the only practical difficulty for global minimization occurring in the uncommon event that we must go throughout nonzero measure "plateaux" of critical points. But the class of quasiconvex functions is not closed by addition. Thus the augmented auxiliary function to be minimized in (1.1) may be nonquasiconvex and indeed multiple isolated local minimizers may occur. Therefore in our opinion the local stationary iteration (1.2) makes much more sense that the previously considered (1.1) for dealing with nonconvex problems.

Under the assumption that f is proper lower semicontinuous and bounded from below on $\mathbb{R}^n_+$ and using a class of proximal distance we will prove that $\{x^k\}$ is well defined and if, in addition, f is quasiconvex it will be proved that $\{f(x^k)\}$ is decreasing and $\{x^k\}$ converges to some point of $U_+ := \{x \in \mathbb{R}^n_+ : f(x) \leqslant \inf_{j\geqslant 0} f(x^j)\}$, assumed nonempty. If $\{\lambda_k\}$ is bounded from above, f is continuously differentiable and d is separable we prove that $\{x^k\}$ converges to a KKT point. Furthermore, we introduce a condition on the proximal distance to obtain the convergence to an optimal point.

The paper is organized as follows: In Section 2 we give some results on regular and general subgradients, quasiconvex theory and a sufficient condition for quasiconvex minimization. In Section 3 we introduce the class of proximal distances that we will use along the paper. In Section 4 we present an algorithm for solving minimization problems with quasiconvex functions on the nonnegative orthant and analyze its convergence properties. Finally, in Section 5 we give our conclusions.

2. Basic results

Throughout this paper $\mathbb{R}^n$ is the Euclidean space endowed with the canonical inner product $\langle , \rangle$ and the norm of x given by $\|x\| := \langle x,x\rangle^{1/2}$. Given an extended real valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ we denote its domain by $\mathtt{dom}f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$. f is said to be proper if $\mathtt{dom}f \neq \emptyset$ and $\forall x \in \mathtt{dom}f$ we have $f(x) > -\infty$.

Finally, f is a lower semicontinuous function if for each $x \in \mathbb{R}^n$ we have that all $\{x^l\}$ such that $\lim_{l\to+\infty} x^l = x$ implies that $f(x) \leqslant \liminf_{l\to+\infty} f(x^l)$. It is easy to prove that the lower semicontinuity of f is equivalent to the closedness of the level set $L_f(\alpha) = \{x \in \mathbb{R}^n : f(x) \leqslant \alpha\}$, for each $\alpha \in \mathbb{R}$.

Definition 2.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper function.

(a) For each $x \in \mathtt{dom}f$, the set of *regular subgradients* (also called *Fréchet subdifferential*) of f at x, denoted by $\hat{\partial}f(x)$, is the set of vectors $s \in \mathbb{R}^n$ such that $\liminf_{y\neq x, y\to x} \frac{1}{\|x-y\|}[f(y) - f(x) - \langle s, y-x\rangle] \geqslant 0$. If $x \notin \mathtt{dom}f$ then $\hat{\partial}f(x) = \emptyset$.
(b) The set of *general subgradients* (also called *limiting subdifferential*) of f at $x \in \mathbb{R}^n$, denoted by $\partial f(x)$, is defined as follows: $\partial f(x) := \{s \in \mathbb{R}^n : \exists x^l \to x, f(x^l) \to f(x), s^l \in \hat{\partial}f(x^l) \text{ and } s^l \to s\}$.
(c) The set of *horizon subgradients* of f at $x \in \mathbb{R}^n$, denoted by $\partial^\infty f(x)$, is defined as follows: $\partial^\infty f(x) := \{v \in \mathbb{R}^n : \exists x^l \to x, f(x^l) \to f(x), v^l \in \hat{\partial}f(x^l) \text{ and } \lambda^l v^l \to v$, for some sequence $\lambda^l \to 0$, with $\lambda^l > 0\}$.

Proposition 2.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper function. Then, the following properties are true*

a. *$\hat{\partial}f(x) \subset \partial f(x)$, for all $x \in \mathbb{R}^n$.*
b. *If f is differentiable at $\bar{x}$ then $\hat{\partial}f(\bar{x}) = \{\nabla f(\bar{x})\}$, so $\nabla f(\bar{x}) \in \partial f(\bar{x})$.*
c. *If f is continuously differentiable in a neighborhood of x, then $\hat{\partial}f(x) = \partial f(x) = \{\nabla f(x)\}$.*
d. *If $g = f + h$ with f finite at $\bar{x}$ and h is continuously differentiable on a neighborhood of $\bar{x}$ then $\hat{\partial}g(\bar{x}) = \hat{\partial}f(\bar{x}) + \nabla h(\bar{x})$ and $\partial g(\bar{x}) = \partial f(\bar{x}) + \nabla h(\bar{x})$.*

Proof. See Rockafellar and Wets (1998, Exercise 8.8.) □

Definition 2.2. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper function. f is called quasiconvex if for all $x, y \in \mathbb{R}^n$, and for all $t \in [0,1]$, it holds that $f(tx + (1-t)y) \leqslant \max\{f(x), f(y)\}$.

Observe that if f is a quasiconvex function then $\mathtt{dom}f$ is a convex set. On the other hand, while a convex function can be characterized by the convexity of its epigraph, a quasiconvex function can be characterized by the convexity of the level sets:

Theorem 2.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper function. Then, f is quasiconvex if and only if the set $\{x \in \mathbb{R}^n : f(x) \leqslant c\}$ is convex for each $c \in \mathbb{R}$.*

Proof. See Bazaara et al. (1993, Theorem 3.5.2). □

Theorem 2.2. *If a proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ has a local minimum at $\bar{x}$ then $0 \in \hat{\partial}f(\bar{x})$.*

Proof. See Rockafellar and Wets (1998, Theorem 10.1). □

Theorem 2.3. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous and quasiconvex function. If $g \in \hat{\partial}f(x)$ and $f(y) \leqslant f(x)$, then $\langle g, y - x\rangle \leqslant 0$.*

Proof. Let $t \in (0,1]$, then from the quasiconvexity of f and the assumption that $f(y) \leqslant f(x)$ we have

$$f(ty + (1-t)x) \leqslant \max\{f(x), f(y)\} = f(x). \tag{2.3}$$

As $g \in \hat{\partial}f(x)$ we obtain

$$f(ty + (1-t)x) \geqslant f(x) + t\langle g, y - x\rangle + o(t), \tag{2.4}$$

where $\lim_{t\to 0} \frac{o(t)}{t} = 0$. From (2.3) and (2.4) we conclude that $t\langle g, y - x\rangle + o(t) \leqslant 0$. Dividing by t and taking $t \to 0$ we obtain the desired result. □

Now, consider the problem

$$\min\{f(x) : g(x) \leqslant 0,\ x \geqslant 0\}, \tag{2.5}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ such that $g(x) = (g_1(x), \ldots, g_m(x))$ with $g_j : \mathbb{R}^n \to \mathbb{R}$, for each $j = 1, \ldots, m$ and $x \geqslant 0$ mean that $x_i \geqslant 0$, for each $i = 1, \ldots, n$.

Theorem 2.4. *Let $f, g_j : \mathbb{R}^n \to \mathbb{R},\ j = 1, \ldots, m,$ be differentiable quasiconvex functions. If $\bar{x}$ satisfies the KKT necessary condition and $\frac{\partial f}{\partial x_i}(\bar{x}) > 0$ for at least one variable x_i ($i \in \{1, 2, \ldots, n\}$), then $\bar{x}$ is a global minimum of the problem (2.5).*

Proof. See Arrow and Enthoven (1961, Theorem 1). □

3. Proximal distance

In this subsection we present the definition of the proximal distance and induced proximal distance, introduced by Auslender and Teboulle (2006), but adapted to our situation for the set $\mathbb{R}^n_+$ which is the constraint of the problem (4.7).

Definition 3.1. A function $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ is called a proximal distance in $\mathbb{R}^n_{++}$ if for each $y \in \mathbb{R}^n_{++}$ it satisfies the following properties:

- **i.** $d(.,y)$ is proper, lower semicontinuous, strictly convex and continuously differentiable on $\mathbb{R}^n_{++}$;
- **ii.** $\mathrm{dom}\, d(.,y) \subset \mathbb{R}^n_+$ and $\mathrm{dom}\,\partial_1 d(.,y) = \mathbb{R}^n_{++}$, where $\partial_1 d(.,y)$ denotes the classical subgradient map of the function $d(.,y)$ with respect to the first variable;
- **iii.** $d(.,y)$ is coercive on $\mathbb{R}^n$ (i.e., $\lim_{\|u\| \to \infty} d(u,y) = +\infty$);
- **iv.** $d(y,y) = 0$.

We denote by $D(\mathbb{R}^n_{++})$ the family of functions satisfying this definition.

Property **i** is needed to preserve convexity of $d(.,y)$, property **ii** will force the iteration of the proximal method to stay in $\mathbb{R}^n_{++}$, and **iii** will be used to guarantee the existence of the proximal iterations. For each $y \in \mathbb{R}^n_{++}$, let $\nabla_1 d(.,y)$ denote the gradient map of the function $d(.,y)$ with respect to the first variable. Note that by definition $d(.,.) \geqslant 0$ and from **iv** the global minimum of $d(.,y)$ is obtained at y, which shows that $\nabla_1 d(y,y) = 0$.

Definition 3.2. Given $d \in D(\mathbb{R}^n_{++})$, a function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ is called the induced proximal distance to d if H is a finite-valued function on $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$ and for each $a, b \in \mathbb{R}^n_{++}$ satisfies

- (**Ii**) $H(a,a) = 0$.
- (**Iii**) $\langle c - b, \nabla_1 d(b,a)\rangle \leqslant H(c,a) - H(c,b),\ \forall c \in \mathbb{R}^n_{++}$.

Let us denoted by $(d,H) \in \mathcal{F}(\mathbb{R}^n_{++})$ to the proximal and induced proximal distance that satisfies the conditions of Definition 3.2.

We also denote $(d,H) \in \mathcal{F}(\mathbb{R}^n_+)$ if there exists H such that:

- (**Iiii**) H is finite valued on $\mathbb{R}^n_+ \times \mathbb{R}^n_{++}$ satisfying (**Ii**) and (**Iii**), for each $c \in \mathbb{R}^n_+$.
- (**Iiv**) For each $c \in \mathbb{R}^n_+, H(c,\cdot)$ has level bounded sets on $\mathbb{R}^n_{++}$.

Finally, we write $(d,H) \in \mathcal{F}_+(\mathbb{R}^n_+)$ if

- (**Iv**) $(d,H) \in \mathcal{F}(\mathbb{R}^n_+)$.
- (**Ivi**) $\forall y \in \mathbb{R}^n_+$ and $\forall \{y^k\} \subset \mathbb{R}^n_{++}$ bounded with $\lim_{k \to +\infty} H(y,y^k) = 0$, then $\lim_{k \to +\infty} y^k = y$.
- (**Ivii**) $\forall y \in \mathbb{R}^n_+$, and $\forall \{y^k\} \subset \mathbb{R}^n_{++}$ such that $\lim_{k \to +\infty} y^k = y$, then $\lim_{k \to +\infty} H(y,y^k) = 0$.

The main result on proximal point method will be when $(d,H) \in \mathcal{F}_+(\mathbb{R}^n_+)$.

Several examples of proximal distances which satisfy the above definitions, for example Bregman distances, proximal distances based on φ-divergences, self-proximal distances, and distances based on second order homogeneous proximal distances, were given by Auslender and Teboulle (2006), Section 3. However, for a good understanding of the reader we give a model example of such proximal distance and we will verify on it that the hypotheses are satisfied.

Example 3.1 (*Bregman distances*). Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper function and differentiable on $int(\mathrm{dom} h)$. We define the function $D_h(.,.) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$D_h(x,y) := h(x) - h(y) - \langle \nabla h(y), y - x\rangle. \tag{3.6}$$

The function h is called a Bregman function with zone S if:

- (**B1**) h is strictly convex and continuous on $\bar{S}$;
- (**B2**) h is continuously differentiable on S;
- (**B3**) Given any $x \in \bar{S}$ and $\delta \in \mathbb{R}$, the right partial level set $L_{D_h}(x,\delta) = \{y \in S : D_h(x,y) \leqslant \delta\}$ is bounded;
- (**B4**) If $\{y^k\} \subset S$ converges to y then $D_h(y,y^k)$ converges to 0.

  When h is a Bregman function then D_h is called the Bregman distance associated with it.

  The classical definition of the Bregman function requires two additional conditions (see for example Eckstein (1993)):
- (**B5**) The left partial level set $L_{D_h}(\delta,y) = \{x \in \bar{S} : D_h(x,y) \leqslant \delta\}$ is bounded for all $y \in S$.
- (**B6**) If $\{z^k\} \subset \bar{S}$ is bounded, $\{y^k\} \subset S$ converges to y, and $\lim_{k \to +\infty} D_h(z^k,y^k) = 0$, then $\lim_{k \to +\infty} z^k = y$.

However, it has been shown that these conditions are redundant as they are implied by conditions (**B**1)–(**B**4); see Bauschke and Borwein (1997).

Let D_h be a Bregman distance, then from (3.6) and (**B**1) it is immediate that for all $x \in \bar{S}$ and $y \in S$, $D_h(x,y) \geqslant 0$ and $D_h(x,y) = 0$ if and only if $x = y$. This is an important property that characterizes D_h as a distance-like function.

Let $\theta : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a proper convex and lower semicontinuous with $(0,+\infty) \subset \mathrm{dom}\ \theta \subset [0,+\infty)$ and such that $\theta \in C^2((0,+\infty))$, $\theta''(t) > 0$, $\forall t > 0$ and $\lim_{t \to 0^+} \theta'(t) = -\infty$. Denote by Θ_0 the class of functions which satisfy the above properties and $\theta(0) < +\infty$ and define $\bar{h}(x) = \sum_{i=1}^n \theta(x_i)$, so

$$D_{\bar{h}}(x,y) = \sum_{i=1}^n (\theta(x_i) - \theta(y_i) - \theta'(y_i)(x_i - y_i)),$$

that is, D_h is separable. It is well known that the class Θ_0 satisfies the conditions (**B**1)–(**B**4) and therefore (**B**5) and (**B**6). Examples of functions in Θ_0 are the following:

(a) $\theta_1(t) = t\log t$ which gives

$$D_{\bar{h}_1}(x,y) = \sum_{i=1}^n \left(x_i \log\frac{x_i}{y_i} + y_i - x_i \right).$$

This function is called the Kullback–Leibler divergence distance.

(b) $\theta_2(t) = t^p - t^q$, if $t \geqslant 0$ and $\theta_2(t) = +\infty$, if $t < 0$; with $p > 1$ and $0 < q < 1$. For $p = 2,\ q = \frac{1}{2}$, we get

$$D_h(x,y) = \|x - y\|^2 + \sum_{i=1}^n \frac{1}{2\sqrt{y_i}} (\sqrt{x_i} - \sqrt{y_i})^2.$$

(c) $\theta_3(t) = \frac{1}{1-p}(pt - t^p)$, if $t \geqslant 0$ and $\theta_3(t) = +\infty$, if $t < 0$; with $p \in (0,1)$, which gives

$$D_h(x,y)=\sum_{i=1}^{n}\left(y_i^p+\frac{1}{(1-p)}(py_i^{p-1}x_i-x_i^p)\right).$$

An interesting particular case is for $p=\frac{1}{2}$, which gives the Bregman distance

$$D_h(x,y)=\sum_{i=1}^{n}\frac{\left(\sqrt{x_i}-\sqrt{y_i}\right)^2}{\sqrt{y_i}}.$$

It is easy to prove that the Bregman separable distance $D_{\bar{h}}$ satisfies the conditions of Definition 3.1 for each $\theta\in\Theta_0$, therefore $D_{\bar{h}}\in D(\mathbb{R}^n_{++})$.

Now define $H(x,y):=D_{\bar{h}}(x,y)$, so $H(a,a)=0,\ \forall a\in\mathbb{R}^n_{++}$ and from the three-points identity, see Lema 3.1 of Chen and Teboulle (1993), we obtain

$$\langle c-b,\nabla_1 D_{\bar{h}}(b,a)\rangle\leqslant H(c,a)-H(c,b).$$

Also, as $D_{\bar{h}}(c,.)$ has level bounded sets on $\mathbb{R}^n_{++}$ we obtain that $(D_{\bar{h}},H)\in\mathcal{F}(\mathbb{R}^n_{++})$. Finally, from (**B6**) and (**B4**) we have respectively

(i) $\forall y\in\mathbb{R}^n_+$ and $\forall\{y^k\}\subset\mathbb{R}^n_{++}$ bounded with $\lim_{k\to+\infty}H(y,y^k)=D_{\bar{h}}(y,y^k)=0$, then $\lim_{k\to+\infty}y^k=y$.
(ii) $\forall y\in\mathbb{R}^n_+$, and $\forall\{y^k\}\subset\mathbb{R}^n_{++}$ such that $\lim_{k\to+\infty}y^k=y$, then $\lim_{k\to+\infty}H(y,y^k)=D_{\bar{h}}(y,y^k)=0$.

Thus, $(D_{\bar{h}},H)\in\mathcal{F}_+(\mathbb{R}^n_+)$.

4. Proximal method

We are interested in solving the problem

$$\min\{f(x):x\geqslant 0\} \tag{4.7}$$

where $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is a proper lower semicontinuous quasiconvex function such that $\mathtt{dom}f\cap\mathbb{R}^n_+\neq\emptyset$ and $x\geqslant 0$ means that each component of x, x_i, is nonnegative.

4.1. Examples

In this subsection we give some examples of minimization problems with quasiconvex objective functions on the nonnegative orthant.

Example 4.1 (*Fractional programming*). The fractional programming problem on $\mathbb{R}^n_+$ is given by $\min\left\{f(x)=\frac{h(x)}{g(x)}:x\geqslant 0\right\}$, where $h,g:\mathbb{R}^n_+\to\mathbb{R}$ and $g(x)>0,\forall x\in\mathbb{R}^n_+$. If the following conditions are satisfied:

(a) h is convex and g is concave
(b) either g is affine or h is nonnegative

then, it is easy to prove that f is a quasiconvex function on $\mathbb{R}^n_+$.

Example 4.2 (*Min–Max Location Problem*). Let $D=\{d_1,d_2,\dots,d_p\}\subset\mathbb{R}^n_+$, $n\geqslant 2$, denote a set of p different demand points and let $x\in\mathbb{R}^n_+$ be the location of a facility to be chosen. If C_i, $i=1,\dots,p$ are compact convex sets with $0\in int(C_i)$, and $int(C_i)$ denoting the interior of C_i, we define the distance between x and d_i by $\gamma_{C_i}(x-d_i)$, with γ_{C_i} the gauge or Minkowsky functional of the set C_i, i.e., $\gamma_{C_i}(x)=\inf\{t>0:x\in tC_i\}$. Observe that if C_i is the unit ball in $\mathbb{R}^n$, then $\gamma_{C_i}(x)$ is the Euclidean distance from x to 0.

To introduce the model, let the function $\gamma:\mathbb{R}^n_+\to\mathbb{R}^p_+$ be given by $\gamma(x)=(\gamma_{C_1}(x-d_1),\dots,\gamma_{C_p}(x-d_p))$, and suppose that the functions $f_i:\mathbb{R}^p\to\mathbb{R}_+,\ i=1,\dots,p$, are nondecreasing on $\mathbb{R}^p_+$, that is, if $x,y\in\mathbb{R}^p_+$ satisfying for each $i=1,\dots,p$, $x_i\leqslant y_i$, then $f_i(x)\leqslant f_i(y)$.

The location model is given by

$$\min\{\phi(x):x\in\mathbb{R}^n_+\}$$

where $\phi(x)=\max_{1\leqslant i\leqslant p}\{\phi_i(x)\}$ with $\phi_i:\mathbb{R}^n_+\to\mathbb{R}$ defined by $\phi_i(x)=f_i(\gamma(x))$, for each $i=1,\dots,p$. This problem arise for example in the location of an emergency unit.

If the functions $f_i:\mathbb{R}^p\to\mathbb{R}_+$ are quasiconvex on $\mathbb{R}^p_+$, then it can be proved that the function ϕ is quasiconvex on $\mathbb{R}^n_+$.

Example 4.3 (*Demand Theory*). In the classical approach to consumer demand, the analysis of consumer behavior begins specifying the consumer's preference $\succeq$ (an " at-least-as-good-as" binary relation) over the commodity bundles in the consumption set $X\subset\mathbb{R}^n_+$. This preference $\succeq$ is assumed to be rational, that is, $\succeq$ is complete and transitive (see Definition 3.B.1 of Mas-Colell et al. (1995)).

A function $\mu:X\to\mathbb{R}$ is said to be an utility function representing a preference relation $\succeq$ on X if the following condition is satisfied

$$x\succeq y,\text{ if and only if }\mu(x)\geqslant\mu(y),$$

for all x, $y\in X$. We should observe that the utility function not always exist. In fact, define in $X=\mathbb{R}^2$ a lexicographic relation:

For $x,y\in\mathbb{R}^2, x\succeq y$ if and only if $x_1>y_1$ or $(x_1=y_1$ and $x_2\geqslant y_2)$.

Fortunately, a very general class of preference relations can be represented by utility functions, see for example Proposition 3.C.1 of Mas-Colell et al. (1995).

If a preference relation $\succeq$ is represented by an utility function μ, then the problem of maximizer the preference of the consumer on X is equivalent to solve the optimization problem

$$(P)\ \max\{\mu(x):x\in X\}.$$

On the other hand, a natural psychological assumption in economy is that the consumer tends to diversify his consumption among all goods, that is, the preference $\succeq$ satisfies the following convexity property: X is convex and if $x\succeq z$ and $y\succeq z$ then $\lambda x+(1-\lambda)y\succeq z$, $\forall\lambda\in[0,1]$.

It can be proved that the if there exists an utility function representing the preference, then the convexity property of $\succeq$ is equivalent to the quasiconcavity of the utility function μ. Thus, (P) becomes in a maximization problem with quasiconcave objective function.

If $X=\mathbb{R}^n_+$ and taking $f=-\mu$ we obtain a minimization problem with a quasiconvex objective function on the nonnegative orthant.

4.2. The algorithm

Now, we propose an extension of the proximal point method with a general proximal distance to solve the problem (4.7). It is as follows:

PPM Algorithm

Initialization: Let $\{\lambda_k\}$ be a sequence of positive parameters and an initial point

$$x^0\in\mathtt{dom}f\cap\mathbb{R}^n_{++}. \tag{4.8}$$

Main Steps: For $k=1,2,3,\dots$ and given $x^{k-1}\in\mathbb{R}^n_{++}$, if $0\in\hat{\partial}f(x^{k-1})$, then stop.
Otherwise, find $x^k\in\mathbb{R}^n_+$ such that

$$0\in\hat{\partial}(f(.)+\lambda_k d(.,x^{k-1}))(x^k), \tag{4.9}$$

where d is a proximal distance such that $(d,H)\in\mathcal{F}_+(\mathbb{R}^n_+)$.

Take $k = k + 1$.

Remark 4.1. Observe that if f is convex then, the iteration (4.9) becomes $x^k = \arg\min_{x\geqslant 0}\{f(x) + \lambda_k d(x,x^{k-1})\}$.

Remark 4.2. As we are interested in solving (4.7) when f is nonconvex it is important to observe (differently from previous research works) that the algorithm only needs, in each iteration, to find a stationary point (and not a global minimum) of the regularized function $f(.) + \lambda_k d(.,x^{k-1})$ so we believe that local algorithms can be used satisfactorily in each iteration.

4.3. Convergence results

Along the paper we assume on f the following Constraint Qualification Condition (CQC):

For all $x\in \mathrm{dom} f\cap\mathbb{R}^n_+$, the only combination of vectors $v_1\in\partial^\infty f(x)$ and $v_2\in\mathcal{N}_{\mathbb{R}^n_+}(x)$ with $v_1 + v_2 = 0$ is $v_1 = v_2 = 0$, where $\mathcal{N}_{\mathbb{R}^n_+}(x)$ is the normal cone of $\mathbb{R}^n_+$.

Remark 4.3. Observe that the above condition holds if f is convex and if $\mathrm{dom} f\cap\mathbb{R}^n_{++}\neq\emptyset$. Of fact, suppose that $v_1\neq 0$, then from the (CQC) and from Rockafellar and Wets (1998, Theorem 8.12) we obtain that $v_1^T(y-x)\leqslant 0,\ \forall y\in\mathbb{R}^n_+$ and $v_1^T(z-x)\geqslant 0,\ \forall z\in\mathrm{dom} f$. That is, $\mathrm{dom} f$ and $\mathbb{R}^n_+$ is separated by the hyperplane $L := \{w\in\mathbb{R}^n : v_1^T(w-x) = 0\}$, which contradicts the assumption $\mathrm{dom} f\cap\mathbb{R}^n_{++}\neq\emptyset$.

Theorem 4.1. *If f is a proper lower semicontinuous function and bounded from below on $\mathrm{dom}\cap\mathbb{R}^n_+$ satisfying (CQC) and $d\in D(\mathbb{R}^n_{++})$, then the sequence $\{x^k\}$, generated by the proximal method (4.8) and (4.9) is well defined and, furthermore, $x^k\in\mathbb{R}^n_{++}$.*

Proof. We proceed by induction. It holds for $k = 0$ due to (4.8). Assume that x^{k-1} exists and $x^{k-1}\in\mathbb{R}^n_{++}$. Define $g(x) = f(x) + \lambda_k d(x,x^{k-1}) + \delta_{\mathbb{R}^n_+}(x)$, where $\delta_{\mathbb{R}^n_+}(.)$ is the indicator function of $\mathbb{R}^n_+$. Then from Definition 3.1, **ii**, we have that $\min\{f(x)+\lambda_k d(x,x^{k-1}) : x\in\mathbb{R}^n_+\}$ is equivalent to $\min\{g(x) : x\in\mathbb{R}^n\}$. Now, due to the lower boundedness and lower semicontinuity of f, as also, the lower semicontinuity and coercivity of $d(.,y)$, there exists $\bar{x}\in\mathbb{R}^n$ (not need unique by the nonconvexity of f) which is a global minimum of g. Thus from Theorem 2.2, $\bar{x}$ satisfies

$$0\in\hat{\partial}\Big(f(.)+\lambda_k d(.,x^{k-1})+\delta_{\mathbb{R}^n_+}(.)\Big)(\bar{x}). \tag{4.10}$$

If $\mathrm{dom} d(.,x^{k-1}) = \mathbb{R}^n_{++}$, then it is immediate that $\bar{x}\in\mathbb{R}^n_{++}$ and from Proposition 2.1, **d**, the iteration (4.9) is obtained from (4.10) for $x^k = \bar{x}$.

Now, suppose that $\mathrm{dom} d(.,x^{k-1}) = \mathbb{R}^n_+$, then from Proposition 2.1, **a**, we have

$$0\in\partial\Big(f(.)+\lambda_k d(.,x^{k-1})+\delta_{\mathbb{R}^n_+}(.)\Big)(\bar{x}). \tag{4.11}$$

Let $f_1 = f$ and $f_2 = \lambda_k d(.,x^{k-1}) + \delta_{\mathbb{R}^n_+}$. As f_2 is convex and $\mathrm{dom} f_2 = \mathbb{R}^n_+$ then from Rockafellar and Wets (1998, Theorem 8.12) we have that $\partial^\infty f_2(x)\subset\mathcal{N}_{\mathbb{R}^n_+}(x)$. Then using the (CQC) we obtain that f_1 and f_2 satisfies the assumptions of Corollary 10.9 of Rockafellar and Wets (1998). Thus,

$$0\in\partial f(\bar{x}) + \lambda_k\partial_1(d(.,x^{k-1}))(\bar{x}) + \mathcal{N}_{\mathbb{R}^n_+}(\bar{x}).$$

From Definition 3.1, **ii**, we conclude that $\bar{x}\in\mathbb{R}^n_{++}$. Finally, taking $x^k = \bar{x}$ and using Proposition 2.1, **d**, we have that (4.10) becomes (4.9). □

Now, we give our assumptions:

Assumption A. f is a proper lower semicontinuous and quasiconvex function.

Assumption B. f is bounded from below on $\mathrm{dom} f\cap\mathbb{R}^n_+$.

As we are interest in the asymptotic convergence of the method, we also assume that in each iteration $0\notin\partial f(x^k)$ which implies that $x^k\neq x^{k-1}$, for all k.

Remark 4.4. From (4.9), Proposition 2.1, **d**, and the smoothness of $\lambda_k d(.,x^{k-1})$, this implies that $0\in\hat{\partial}f(x^k)+\lambda_k\nabla_1 d(x^k,x^{k-1})$. Thus, there exists $g^k\in\hat{\partial}f(x^k)$ such that

$$g^k = -\lambda_k\nabla_1 d(x^k,x^{k-1}).$$

Proposition 4.1. *Under assumptions A, B and $d\in D(\mathbb{R}^n_{++})$, we have that $\{f(x^k)\}$ is decreasing and converges.*

Proof. As $x^k\neq x^{k-1}$ and $d(.,x^{k-1})$ is strictly convex then

$$\langle\nabla_1 d(x^k,x^{k-1}) - \nabla_1 d(x^{k-1},x^{k-1}), x^k - x^{k-1}\rangle > 0 \tag{4.12}$$

As $\nabla_1 d(x^{k-1},x^{k-1}) = 0$, $\lambda_k > 0$ and from Remark 4.4, we obtain $\langle g^k, x^{k-1} - x^k\rangle > 0$. Using the quasiconvexity of f and Theorem 2.3, this implies that $f(x^k) < f(x^{k-1})$. The convergence of $\{f(x^k)\}$ is immediate from the lower boundedness of f. □

Now, we define the following set $U_+ := \{x\in\mathbb{R}^n_+ : f(x)\leqslant \inf_{j\in\mathbb{N}} f(x^j)\}$. Observe that this set depends on the choice of the initial iterates x^0 and the sequence $\{\lambda_k\}$. Furthermore, if $U_+ = \emptyset$ then it is easy to prove that $\lim_{k\to+\infty} f(x^k) = \inf_{x\geqslant 0} f(x)$, and $\{x^k\}$ is unbounded.

From now on we assume that $U_+\neq\emptyset$. From the assumptions we obtain that U_+ is a nonempty, closed and convex set (see Theorem 2.1 for the convexity property).

Proposition 4.2. *Under assumptions A, B and $(d,H)\in\mathcal{F}(\mathbb{R}^n_{++})$, the sequence $\{x^k\}$, generated by the proximal method, (4.8) and (4.9), is H-Fejér convergent to U_+, that is,*

$$H(x,x^k)\leqslant H(x,x^{k-1}),\quad \forall x\in U_+.$$

Proof. Let $x\in U_+$, then $f(x)\leqslant f(x^k)$. From Remark 4.4 and using Theorem 2.3 we obtain $\langle\nabla_1 d(x^k,x^{k-1}), x - x^k\rangle\geqslant 0$. Then, using property **Iii** of Definition 3.2 we have the result. □

Lemma 4.1. *Under assumptions A, B and $(d,H)\in\mathcal{F}(\mathbb{R}^n_+)$, the sequence $\{x^k\}$ is bounded.*

Proof. From the previous proposition we deduce that $H(x,x^k)\leqslant H(x,x^0) = \alpha$, $\forall x\in U_+$. This implies that $x^k\in L_H(x,\alpha) = \{y\in\mathbb{R}^n_+ : H(x,y)\leqslant\alpha\}$, and from Definition 3.2, **Iiv**, $L_H(x,\alpha)$ is bounded. Thus $\{x^k\}$ is bounded. □

Theorem 4.2. *Under assumptions A, B and $(d,H)\in\mathcal{F}_+(\mathbb{R}^n_+)$, the sequence $\{x^k\}$ converges to some point of U_+.*

Proof. From previous lemma $\{x^k\}$ is bounded, then there exists a subsequence $\{x^{k_j}\}$ which converges to $\bar{x}$, that is $\lim_{j\to+\infty} x^{k_j} = \bar{x}$. As f is lower semicontinuous, nonincreasing and $\{f(x^k)\}$ converges we have $\bar{x}\in U_+$. Suppose that there exists another sequence $\{x^{k_l}\}$ such that $\lim_{l\to+\infty} x^{k_l} = x^*\in U_+$. Using property **Ivii**, from Definition 3.2, we obtain $\lim_{l\to+\infty} H(x^*,x^{k_l}) = 0$. Due to $\{H(x^*,x^k)\}$ is non-

increasing and bounded from below then it converges and therefore $\lim_{k\to+\infty} H(x^*, x^k) = 0$. Thus $\lim_{j\to+\infty} H(x^*, x^{k_j}) = 0$. Using property **Ivi**, from Definition 3.2, we obtain that $\lim_{j\to+\infty} x^{k_j} = x^*$, that is, $\bar{x} = x^*$. $\square$

To obtain further information on the limit point we restrict the class of proximal distances given in Definition 3.1 to a special class of separable proximal distances, that is, $d(x,y) = \sum_{i=1}^n d_i(x_i, y_i)$, where each $d_i : \mathbb{R}\times\mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ is a continuously differentiable function on $\mathbb{R}_{++}\times\mathbb{R}_{++}$, such that if there exist i with

$$\frac{\partial d_i}{\partial x_i}(x_i, y_i) > 0 \Rightarrow x_i > y_i. \tag{4.13}$$

Remark 4.5. Observe that the well known Bregman separable, φ-divergence and second order homogeneous distances satisfy the above property.

Theorem 4.3. *Suppose that $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is continuously differentiable and quasiconvex and assumption B and $(d,H) \in \mathcal{F}_+(\mathbb{R}^n_+)$ are satisfied, where d is separable and (4.13) holds. If $0 < \lambda_k < \bar{\lambda}$, for all k and some $\bar{\lambda} > 0$, then the sequence $\{x^k\}$ converges to a KKT point of the problem (4.7).*

Proof. From the previous theorem, there exists $\bar{x} \in U_+$ such that $\lim_{k\to+\infty} x^k = \bar{x}$. We are going to show that $\bar{x}$ is a KKT point, that is,

$$\bar{x}_i \geqslant 0, \quad (\nabla f(\bar{x}))_i \geqslant 0, \quad \bar{x}_i(\nabla f(\bar{x}))_i = 0, \quad \forall i = 1,\ldots,n.$$

Due to Theorem 4.1 the first condition is satisfied. To prove the other conditions consider the following sets $I(\bar{x}) = \{i \in \{1,\ldots,n\} : \bar{x}_i = 0\}$, and $J(\bar{x}) = \{i \in \{1,\ldots,n\} : \bar{x}_i > 0\}$.

We will prove that $(\nabla f(\bar{x}))_i \geqslant 0$. Let $i \in I(\bar{x})$ and suppose that $(\nabla f(\bar{x}))_i < 0$. From continuity of ∇f we have there exists $k_0 \in \mathbb{N}$ such $(\nabla f(x^k))_i < 0$, $\forall k \geqslant k_0$. From Remark 4.4 that implies that $\frac{\partial d_i}{\partial x_i}(x_i^k, x_i^{k-1}) > 0$, $\forall k \geqslant k_0$, and from assumption (4.13) we have

$$x_i^k > x_i^{k-1} > \cdots > x_i^{k_0-1} > 0, \quad \forall k \geqslant k_0.$$

As $x_i^{k_0-1}$ is a fixed point, there exists $\gamma > 0$ such that

$$x_i^k > x_i^{k-1} > \cdots > x_i^{k_0-1} > \gamma > 0, \quad \forall k \geqslant k_0.$$

Taking $k \to +\infty$ we obtain that $0 \geqslant \gamma > 0$, which is a contradiction. Therefore $(\nabla f(\bar{x}))_i \geqslant 0$, for all $i \in I(\bar{x})$. Now let $i \in J(\bar{x})$. From continuity of ∇f and Remark 4.4 we have

$$(\nabla f(\bar{x}))_i = \lim_{k\to\infty} (\nabla f(x^k))_i = \lim_{k\to\infty} -\lambda_k \frac{\partial d_i}{\partial x_i}(x_i^k, x_i^{k-1}). \tag{4.14}$$

As $x_i^{k-1} \to \bar{x}_i > 0$, d_i is continuously differentiable on $\mathbb{R}_{++}\times\mathbb{R}_{++}$ and $\nabla_1 d(x,x) = 0$, $\forall x \in \mathbb{R}^n_{++}$, we have $\lim_{k\to\infty} \frac{\partial d_i}{\partial x_i}(x_i^k, x_i^{k-1}) = \frac{\partial d_i}{\partial x_i}(\bar{x}_i, \bar{x}_i) = 0$. Now, using in (4.14) the boundedness of $\{\lambda_k\}$ and the above result we have $(\nabla f(\bar{x}))_i = 0$, for $i \in J(\bar{x})$.

Finally, $\bar{x}_i(\nabla f(\bar{x}))_i = 0$ is immediate from above results. $\square$

Corollary 4.1. *Under the same assumptions of the previous theorem and the existence of the optimal solutions of (4.7) we have that if there exists $i \in I(\bar{x})$ such that $\lim_{k\to+\infty}\left(\lambda_k \frac{\partial d_i}{\partial x_i}(x_i^k, x_i^{k-1})\right) < 0$, then $\{x^k\}$ converges to a solution of (4.7).*

Proof. Let $\{x^k\}$ such that $\lim_{k\to\infty} x^k = \bar{x}$. From the previous theorem $\bar{x}$ is a KKT point of the problem (4.7). Suppose that $\lim_{k\to+\infty}\left(\lambda_k \frac{\partial d_i}{\partial x_i}(x_i^k, x_i^{k-1})\right) < 0$, for some $i \in I(\bar{x})$. From Remark 4.4 and continuity of ∇f we obtain $\frac{\partial f}{\partial x_i}(\bar{x}) > 0$. Now using item **a** of Theorem 2.4, for the case when $g(x) = -x$, we obtain that $\bar{x}$ is an optimal solution of the problem (4.7). $\square$

Remark 4.6. The same ideas used in the proof of the previous theorem can be used to extend the convergence result of the continuous method studied by Attouch and Teboulle (2004).

5. Conclusion and some discussion

Our approach contributes to the progress in the efficient solution of minimization problems with quasiconvex objective functions using an extension of the proximal point method. Previous results turn out to be particular cases of our general approach as also the full convergence for a large class of proximal distances, which was not studied previously, is obtained. Furthermore, we give a sufficient condition on the proximal distance to obtain the convergence of the method to an optimal solution of the problem.

To the author's knowledge, this is the first attempt to develop a method with general proximal distances to solve minimization problems with quasiconvex objective functions on the nonnegative orthant.

For a computational implementation of the proposed method it is needed to solve the iteration (4.9) using a local minimization algorithm, which only provides an approximate solution. Therefore, we consider that in a future work it is important to analyze the convergence of the proposed algorithm considering now an inexact iteration. Some classical candidates are the following:

(a) Given a sequence of parameters $\{\epsilon_k\} \subset \mathbb{R}_+$ such that $\sum_{i=1}^n \epsilon_k < +\infty$ and $x^{k-1} \in \mathbb{R}^n_{++}$, find $x^k \in \mathbb{R}^n_+$ and $g^k \in \hat{\partial} f(x^k)$ such that

$$\|g^k + \lambda_k \nabla_1 d(x^k, x^{k-1})\| \leqslant \epsilon_k.$$

(b) Given a sequence of vectors $\{e^k\} \subset \mathbb{R}^n$ and $x^{k-1} \in \mathbb{R}^n_{++}$, find $x^k \in \mathbb{R}^n_+$ and $g^k \in \hat{\partial} f(x^k)$ such that

$$g^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = e^k,$$

where $\{e^k\}$ should satisfies some desirable properties.

(c) Find $x^k \in \mathbb{R}^n_+$ and $g^k \in \hat{\partial}_{\epsilon_k} f(x^k)$ such that

$$g^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = 0,$$

where

$$\hat{\partial}_\epsilon f(x) = \left\{ s \in \mathbb{R}^n : \liminf_{y\neq x, y\to x} \frac{1}{\|x-y\|}[f(y) - f(x) - \langle s, y-x\rangle] \geqslant -\epsilon \right\}.$$

Despite that theoretical convergence results of the method for quasiconvex functions using inexact iteration have not been developed, some computational experiments were reported, for differentiable and quasiconvex functions, by Cunha et al. (2010), Chen and Pan (2008), Pan and Chen (2007) and Souza et al. (2010) and their results attest that the inexact iteration *a*), is very promising for finding the global optimal solution of quasiconvex functions on the nonnegative orthant.

On the other hand, Attouch et al. (2010) have been developed a convergence analysis of the inexact proximal algorithm for nonconvex objective functions satisfying the Kurdyka–Lojasiewicz inequality (see Section 4 of that paper) and we conclude that the inexact iteration proposed by these authors may be useful for our approach.

Acknowledgments

We are grateful to the anonymous referees that suggested important inclusions on the paper.

References

Alvarez, F., Bolte, J., Brahic, O., 2004. Hessian riemannian gradient flows in convex programming. SIAM Journal of Optimization 43 (2), 477–501.

Arrow, K.J., Enthoven, A.C., 1961. Quasi-concave programming. Econometrica 29 (4), 779–800.

Attouch, H., Bolte, J., 2009. On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. Mathematical Programming 116 (1-2, Ser. B), 5–16.

Attouch, H., Teboulle, M., 2004. Regularized Lotka–Volterra dynamical system as continuous proximal-like method in optimization. Journal Optimization Theory and Applications 121, 541–580.

Attouch, H., Bolte, J., Svaiter, B., 2010. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Optimization. Online.

Auslender, A., Teboulle, M., 2006. Interior gradient and proximal methods for convex and conic optimization. SIAM Journal of Optimization 16 (3), 697–725.

Auslender, A., Teboulle, M., Ben-Tiba, S., 1999. Interior proximal and multiplier methods based on second order homogeneous functionals. Mathematics of Operations Research 24 (3), 645–668.

Barron, N., Liu, W., 1997. Calculus of variation l^{∞}. Applied Mathematical Optimization 35, 237–263.

Bauschke, H.H., Borwein, J.M., 1997. Legendre functions and the method of random Bregman projections. Journal of Convex Analysis 4, 27–67.

Bazaara, M.S., Sherali, H.D., Shetty, C.M., 1993. Nonlinear Programming: Theory and Algorithms, second ed. John Wiley and Sons, New York, NY.

Chen, J.S., Pan, S., 2008. A proximal-like algorithm for a class of nonconvex programming. Pacific Journal of Optimization 4, 319–333.

Chen, G., Teboulle, M., 1993. Convergence analysis of a proximal-like minimization algorithm using Bregman functions 3, 538–543.

Cunha, F.G.M., da Cruz Neto, J.X., Oliveira, P.R., 2010. A proximal point algorithm with ϕ-divergence to quasiconvex programming 59, 777–792.

Eckstein, J., 1993. Nonlinear proximal point algorithms using Bregman functions with applications to convex programming. Mathematics of Operations Research 18, 206–226.

Goudou, X., Munier, J., 2009. The gradient and heavy ball with friction dynamical system: The quasiconvex case. Mathematical Programming, Series B 116, 173–191.

Gromicho, J., 1998. Quasiconvex Optimization and Location Theory. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Iusem, A., Teboulle, M., 1995. Convergence rate analysis of nonquadratic proximal method for convex and linear programming. Mathematics of Operations Research 20, 657–677.

Iusem, A., Svaiter, B., Teboulle, M., 1994. Entropy-like proximal methods in convex programming. Mathematics of Operations Research 19, 790–814.

Kiwiel, K.C., 1997. Proximal minimization methods with generalized Bregman functions. SIAM Journal of Control and Optimization 35, 1142–11268.

Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, New York, NY, USA.

Pan, S., Chen, J.S., 2007. Entropy-like proximal algorithms based on a second-order homogeneous distances function for quasiconvex programming. Journal of Global Optimization 39, 555–575.

Rockafellar, R.T., Wets, R., 1998. Variational analysis. Grundlehren der Mathematischen, Wissenschaften, vol. 317. Springer.

Souza, S.S., Oliveira, P.R., da Cruz Neto, J.X., Soubeyran, A., 2010. A proximal method with separable Bregman distance for quasiconvex minimization on the nonnegative orthant. European Journal of Operational Research 201, 365–376.

Takayama, A., 1995. Mathematical Economics, 2nd ed. Cambridge Univ. Press.

Teboulle, M., 1992. Entropic proximal mappings with application to nonlinear programming. Mathematics of Operation Research 17, 670–690.

Teboulle, M., 1997. Convergence of proximal-like algorithms. SIAM Journal on Optimization 7, 1069–1083.